

The robustness of magic state distillation against errors in Clifford gates

Tomas Jochym-O'Connor,^{1,*} Yafei Yu,^{1,2} Bassam Helou,¹ and Raymond Laflamme^{1,3}

¹ *Institute for Quantum Computing and Department of Physics and Astronomy,
University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1, Canada*

² *Laboratory of Nanophotonic Functional Materials and Devices,*

SIPSE & LQIT, South China Normal University, Guangzhou 510006, China

³ *Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo Ontario, N2L 2Y5, Canada*

Quantum error correction and fault-tolerance have provided the possibility for large scale quantum computations without a detrimental loss of quantum information. A very natural class of gates for fault-tolerant quantum computation is the Clifford gate set and as such their usefulness for universal quantum computation is of great interest. Clifford group gates augmented by magic state preparation give the possibility of simulating universal quantum computation. However, experimentally one cannot expect to perfectly prepare magic states. Nonetheless, it has been shown that by repeatedly applying operations from the Clifford group and measurements in the Pauli basis, the fidelity of noisy prepared magic states can be increased arbitrarily close to a pure magic state [1]. We investigate the robustness of magic state distillation to perturbations of the initial states to arbitrary locations in the Bloch sphere due to noise. Additionally, we consider a depolarizing noise model on the quantum gates in the decoding section of the distillation protocol and demonstrate its effect on the convergence rate and threshold value. Finally, we establish that faulty magic state distillation is more efficient than fault-tolerance-assisted magic state distillation at low error rates due to the large overhead in the number of quantum gates and qubits required in a fault-tolerance architecture. The ability to perform magic state distillation with noisy gates leads us to conclude that this could be a realistic scheme for future small-scale quantum computing devices as fault-tolerance need only be used in the final steps of the protocol.

PACS numbers: 03.67.Pp, 03.67.Ac, 03.67.Lx

Keywords: Universal quantum computation, magic state distillation, quantum fault-tolerance

I. INTRODUCTION

Processes such as imperfect control of quantum operations or unintended coupling between qubit systems and their environment lead to errors in any realistic implementation of a quantum computing device. As such, quantum error correction has been developed in order to recover the quantum information that would otherwise be lost due to these faults [2–4]. However, quantum error correction itself is not enough for the implementation of a robust quantum computing device as errors that occur between the error correction steps could propagate between qubits. Propagating errors could prove to be detrimental to the recovery of quantum information and need to be avoided in order to implement any realistic error correction scheme. Fault-tolerant quantum computation aims to address this concern by encoding the information of each qubit into a larger Hilbert space of many qubits and perform encoded quantum gates in such a way that errors do not propagate through multiple qubit blocks [5]. Fault-tolerance allows the faults to remain tractable at the cost of needing a polylogarithmic increase in the number of qubits and quantum gates to perform encoded operations when the error rate of the quantum gates is below a certain target threshold [6–10].

A desired property of encoded gate operations is transversality, which prevents the propagation of errors in the encoded states of the fault-tolerant architecture. The Clifford gate set, the group of gates generated by the Hadamard gate, the phase gate, and CNOT gate, has been shown to be transversal for many quantum codes [11]. The Clifford gates, along with measurement and state preparation in the Z -eigenbasis form the class of stabilizer operations, which have been shown to be efficiently classically simulatable [12]. In order to use Clifford gates for universal quantum computation, one requires the ability to produce a pure, single-qubit, non-stabilizer state [13], known as a magic state. Perfect magic state preparation is difficult due to experimental errors, however magic state distillation allows for the creation of an arbitrarily high fidelity magic state from noisy ancillas by repeatedly applying stabilizer operations [1, 14, 15]. In this work we investigate the effect of noise, in the state preparation and quantum gate application, on the convergence of the magic state distillation protocol.

*Electronic address: trjochym@uwaterloo.ca

In Bravyi and Kitaev's magic state distillation protocol [1], five copies of a noisy magic state are used to extract a single state of higher fidelity with respect to the magic state. The process is then iterated to increase the fidelity to be arbitrarily high. The input states to the distillation scheme are assumed to be along the magic state axis in the Bloch sphere, the axis connecting the magic state and its orthogonal complement. This assumption is based upon the ability to perform a dephasing channel that collapses all points of the Bloch sphere to their projection along this axis. In this work, we extend the analysis of the distillation scheme to one where the location of the input state is an arbitrary state in the Bloch sphere. The motivation for this is twofold, performing a dephasing operation in an experimental setup would introduce errors that would manifest themselves as perturbations off the magic state axis; moreover, there may be states off the magic state axis that converge faster to the magic state after multiple iterations than the states with the same fidelity on the magic state axis. Performing such an analysis will enable us to conclude that the magic state distillation protocol is robust to slight perturbations about the magic state axis in the Bloch sphere due to noise.

In Section III, we turn our attention to the effect of noise present in the quantum gates of the distillation protocol. We investigate the consequences of depolarizing noise, a one-parameter noise model, on the one and two-qubit Clifford gates in the five-qubit decoder of the protocol. Such noise will affect the rate of convergence, increase the threshold for the input states of the protocol, and pose a restriction on the ability to prepare a magic state with arbitrarily high fidelity. As such, with the development of fault-tolerant schemes for the reduction of noise, one can ask if it is more efficient in the total number of quantum gates to use faulty gates or fault-tolerant encoded gates to perform the distillation protocol. In Section IIIB we show that using faulty Clifford gates is the more efficient scheme and using a fault-tolerant encoding is unnecessary unless the required gate fidelity of the universal gate is much higher than that of the Clifford gates at one's disposal. The ability to perform magic state distillation without the use of a fault-tolerant encoding is promising for future experimental realizations of multiple rounds of state distillation on small-scale quantum computers.

II. THE EVOLUTION OF QUANTUM STATES UNDER MAGIC STATE DISTILLATION WITH PERFECT CLIFFORD GATES

The input states to Bravyi and Kitaev's magic state distillation protocol are assumed to be along the magic state axis [1]. However, preparing such states may prove to be difficult experimentally. In this section we study the convergence of the distillation scheme under perturbations about the magic state axis. Moreover, we show that for low fidelity input states, the convergence to the magic state may be improved for states away from the magic state axis. This suggests that while performing a dephasing operation to initialize the input states to be along the axis may be useful, it is not absolutely necessary in certain fidelity regimes.

The Clifford gate set is generated by the following gates,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad CNOT = \begin{pmatrix} I & 0 \\ 0 & \sigma^x \end{pmatrix}, \quad (1)$$

where each individual gate can be performed on any qubit, and CNOT can be performed on any pair of qubits where I and σ^x are 2-by-2 Pauli matrices. The power of magic state distillation is that it requires only the use of Clifford gate operations, along with $|0\rangle$ state ancilla preparation and measurement in the Z -basis, to distill multiple copies of noisy magic states to one of higher fidelity, provided the initial state is above a given threshold. This is appealing as Clifford gates have been shown to have transversality in many quantum error correcting codes [11], and as such can be implemented fault-tolerantly in order to reduce their error rate.

Let σ^i denote the i^{th} Pauli matrix in the computational basis. There are two types of magic states, up to one-qubit Clifford operators, the H-type magic state,

$$|H\rangle\langle H| = \frac{1}{2} \left[I + \frac{1}{\sqrt{2}}(\sigma^x + \sigma^z) \right],$$

which can be used to implement the $\pi/8$ -phase gate, $\Lambda_{\pi/8} = e^{i\pi/8}|0\rangle\langle 0| + e^{-i\pi/8}|1\rangle\langle 1|$, and the T-type magic state,

$$|T\rangle\langle T| = \frac{1}{2} \left[I + \frac{1}{\sqrt{3}}(\sigma^x + \sigma^y + \sigma^z) \right],$$

which can be used to implement the $\pi/12$ -phase gate $\Lambda_{\pi/12}$ [1], both of which, along with the Clifford gates, provide universal quantum computation. Many efforts have contributed to building protocols for magic state distillation and achieving tight noise thresholds for the noisy input ancillas to the distillation protocol in order to understand the

given by coordinates (x, y, z) , then the output state coordinates are as follows:

$$\begin{aligned} x_{out} &= \frac{-z(z^4 - 5x^2 + 5y^2(x^2 - 1))}{(1 + 5z^2y^2 + 5x^2y^2 + 5z^2x^2)}, \\ y_{out} &= \frac{-y(y^4 - 5z^2 + 5x^2(z^2 - 1))}{(1 + 5z^2y^2 + 5x^2y^2 + 5z^2x^2)}, \\ z_{out} &= \frac{-x(x^4 - 5z^2 + 5y^2(z^2 - 1))}{(1 + 5z^2y^2 + 5x^2y^2 + 5z^2x^2)}, \end{aligned} \quad (5)$$

The plane of states that have fidelity F are the states in the Bloch sphere that satisfy,

$$x + y + z = \sqrt{3}(2F - 1). \quad (6)$$

For a given plane of constant input fidelity F_{in} , define a new coordinate system for that plane where r is the radial distance of the input state from the magic state axis, and θ as the angle between the distance vector and the modified x axis, as shown in Figure 2. The difference between the input fidelity and output fidelity can then be expressed according to these coordinates as

$$d = F_{out} - F_{in} = -\frac{2\left(a(54 - 60a^2 + 14a^4 + 135r^4) + 15\sqrt{6}(-3 + a^2)r^3 \cos 3\theta\right)}{2\sqrt{3}\left(108 + 20a^4 + 135r^4 + 60\sqrt{6}ar^3 \cos 3\theta\right)}, \quad (7)$$

where a is related to the input fidelity, $a = \sqrt{3}(2F_{in} - 1)$. As such, the ability of the distillation protocol to increase the fidelity of the input states depends on the distance of the initial states from the magic state axis and on the spatial angle with respect to the modified x axis as well as the input fidelity.

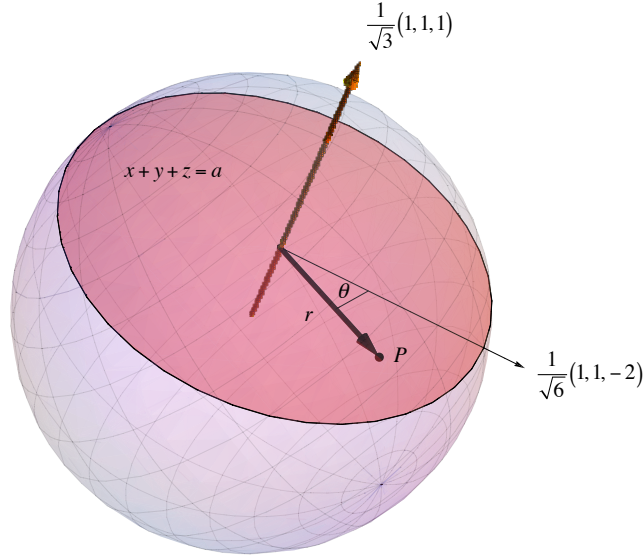


FIG. 2: Labelling of the coordinates in the Bloch sphere used to derived the fidelity difference after a single run of the magic state distillation protocol, given in Eq. 7. The red plane is the plane of constant fidelity $x + y + z = a$, the orange arrow is the magic state axis, labeled by its normalized Bloch sphere coordinates $(1, 1, 1)/\sqrt{3}$, and the axis labelled by the vector $(1, 1, -2)/\sqrt{6}$ is the modified x -axis. For a given point P , the coordinates a , r , and θ are defined as in the figure.

Figure 3 demonstrates the dependence on the distance from the magic state axis and angle for convergence to the magic state after many iterations for the set of states on the plane with fidelity $F = 0.886$ in the positive octant ($x, y, z > 0$). States close to the magic state axis converge to the magic state, while states far away from the axis converge to other undesirable states in the Bloch sphere. Perhaps more surprisingly, there are states below the distillation threshold set by Bravyi and Kitaev [1] that can converge to the magic state. As Figure 4 shows, states away from the magic state axis on the fidelity plane of $F = 0.8270$ converge to the magic state, while those

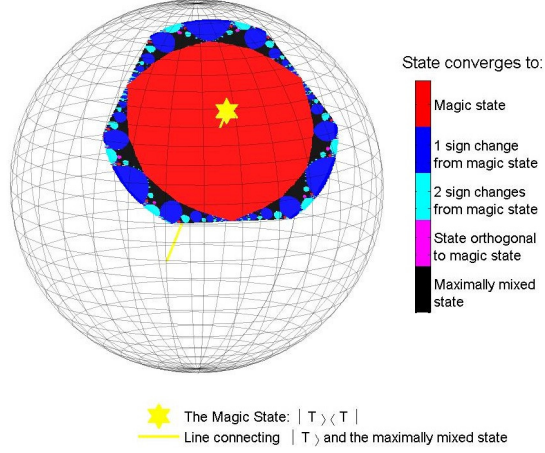


FIG. 3: Final states of convergence after multiple rounds of distillation for states on the fidelity plane $F = 0.886$, that is a cut of the Bloch sphere with states of a fixed fidelity, as given by Eq. 6. The states in red represent the states that converge to the magic state after multiple rounds of distillation. The states in light and dark blue converge to states with the coordinates $(\pm 1, \pm 1, \pm 1)/\sqrt{3}$ in the Bloch sphere, where the number of sign changes from the coordinates of the $|T\rangle$ magic state $(1, 1, 1)/\sqrt{3}$ is given by the shading of blue. The states in pink converge to the state orthogonal to $|T\rangle$, and black to the maximally mixed state.

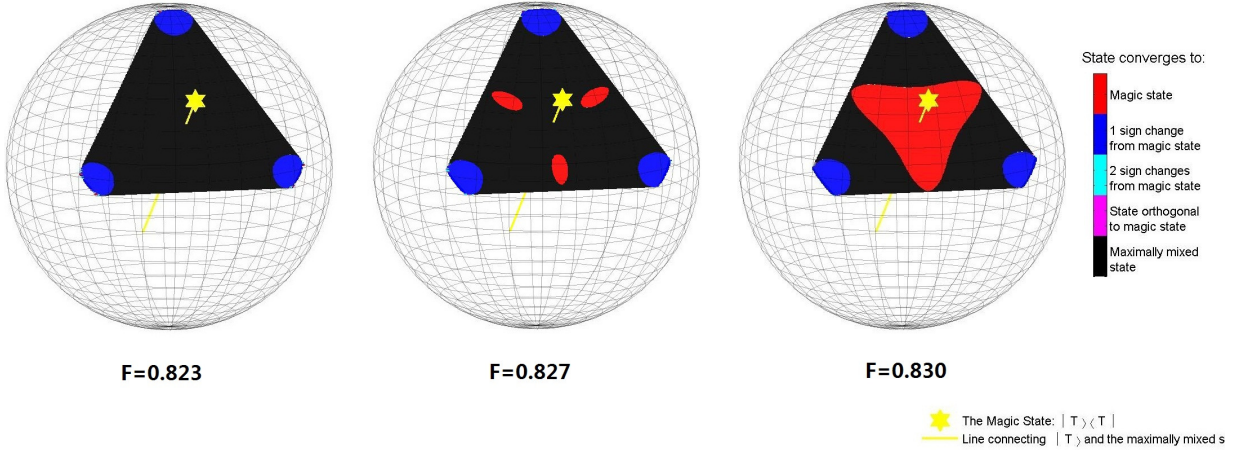


FIG. 4: Convergence of the states around the perfect Clifford gate threshold 0.8273 for states along the magic state axis. All states much below the theoretical threshold ($F = 0.823$) converge away from the magic state axis, yet some states slightly below the threshold for on axis convergence converge to the magic state ($F = 0.8270$). This convergence only appears for states close to the three angles of maximum output fidelity on the fidelity plane. All states close to the magic state axis converge to the magic state for fidelities above the threshold ($F = 0.830$).

on the axis converge to the maximally mixed state, as expected as F is below the fidelity threshold for states on the axis. Notice in Eq. 7 that the maximal increase in the fidelity of the state after the distillation procedure occurs for angles $\theta = 0, 2\pi/3, 4\pi/3$. Fixing $\theta = 0$, the difference in fidelity can show an increase as a function of the distance from the magic state axis r for small distances before dropping off rapidly as the distance from the axis grows. For an input fidelity just below the fidelity threshold set in Ref. [1], the fidelity difference can be negative for states close to the axis and increase to be positive for certain distances away from the axis, as Figure 5(a) shows, but only for states whose angle is close to the angles of maximal increase, as plotted in Figure 5(b). Therefore at

angles close to the angles of maximal increase, one can obtain states whose fidelity threshold can be below the on-axis threshold of $\frac{1}{2}(1 + \sqrt{\frac{3}{7}}) \approx 0.8273$. As such, by performing the magic state distillation routine without the dephasing transformation, one can slightly lower the threshold for the fidelity of input states to be 0.825 for states at the angle of maximal increase.

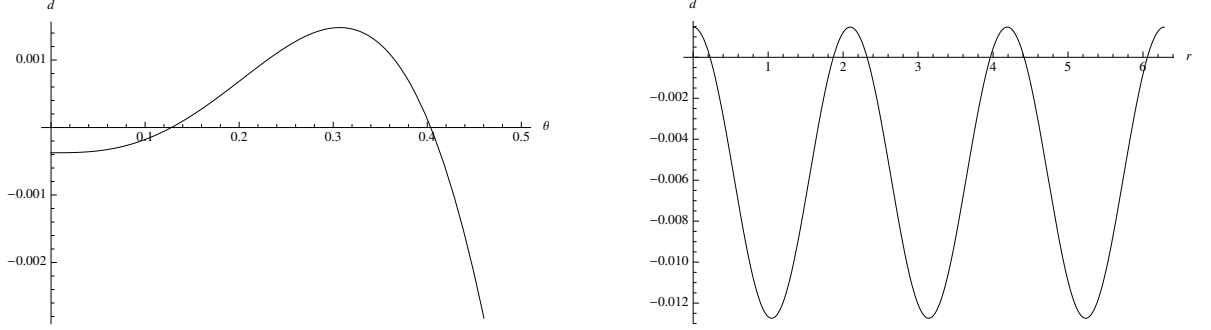


FIG. 5: (a) Dependence of the difference in fidelity d , as defined in Equation (7), between the initial and final states as a function of distance from the magic state axis in the Bloch sphere, at fixed angle $\theta = 0$ and fidelity $F = 0.8269$. (b) Dependence of distance in fidelity between the initial and final states as a function of angle for fixed distance from the Bloch sphere $r = 0.3$ and fidelity $F = 0.8269$.

We may thus conclude that the dephasing operation that projects the noisy magic states onto the magic state axis is not necessarily needed as there are regions of convergence around this axis. As such, noisy dephasing processes that would have slight perturbations off the magic state axis would not be detrimental to the convergence to the magic state. Furthermore, for certain states off the magic state axis and of fidelity just below that derived by Bravyi and Kitaev, omitting the dephasing operation would be beneficial as it would allow such states to remain useful for magic state distillation.

III. MAGIC STATE DISTILLATION WITH NOISY CLIFFORD GATES

A. Error rate of faulty magic state distillation

In an experimental realization of the implementation of magic state distillation, any quantum gate will introduce noise. As such, an interesting experimental question would be: to what level of noise is the distillation still beneficial? In order to address such a question, we consider the case where all the gates in the decoding circuit, shown by the boxed region in Figure 1, are subjected to depolarizing noise. Depolarizing noise is a common noise model for physical implementations of quantum information processing. A noisy one-qubit Clifford gate G will be modelled as follows:

$$\boxed{G} \rightarrow \boxed{G} \boxed{\Lambda_1},$$

where the gate Λ_1 performs the depolarizing transformation with a noise parameter p_1 ,

$$\Lambda_1(\rho, p_1) \rightarrow (1 - p_1)\rho + p_1 \frac{I}{2} = (1 - \frac{3p_1}{4})\rho + \frac{p_1}{4}(\sigma^x \rho \sigma^x + \sigma^y \rho \sigma^y + \sigma^z \rho \sigma^z). \quad (8)$$

Similarly, we introduce noise to a two-qubit controlled Clifford gate by adding a two-qubit depolarizing gate,

$$\begin{array}{c} \bullet \\ | \\ \boxed{G} \end{array} \rightarrow \begin{array}{c} \bullet \\ | \\ \boxed{G} \end{array} \boxed{\Lambda_2},$$

where the two-qubit depolarizing gate Λ_2 is defined by the transformation with a two-qubit noise parameter p_2 ,

$$\begin{aligned} \Lambda_2(\rho, p_2) &\rightarrow (1 - p_2)\rho + p_2 \frac{I_2}{4} \\ &= \left(1 - \frac{15p_2}{16}\right)\rho + \frac{p_2}{16} \left((I \otimes \sigma_x)\rho(I \otimes \sigma_x) + (I \otimes \sigma_y)\rho(I \otimes \sigma_y) + \dots + (\sigma_z \otimes \sigma_y)\rho(\sigma_z \otimes \sigma_y) + (\sigma_z \otimes \sigma_z)\rho(\sigma_z \otimes \sigma_z) \right). \end{aligned} \quad (9)$$

We shall only consider errors affecting the gates in the decoding procedure. In order to compare the fidelity threshold for the noisy decoding procedure with the ideal magic state distillation protocol proposed by Bravyi and Kitaev [1], we assume the input states to the decoding procedure are along the magic state axis. This assumption is valid in the error regime we will be considering, as depolarizing errors in the dephasing transformation would lead to a negligible deviation of the input states from the magic state axis. We can relate the parameter p_i to the error per gate, which is defined as $E_i = 1 - F_i$, where F_i is the fidelity of the gate i . These values are related by the simple relationships: $p_1 = 2E_1$ and $p_2 = 4E_2/3$.

Assuming the error rate of the one and two-qubit gates are low (omitting quadratic and higher order terms in p_1 and p_2), the output (unnormalized) state will have the following matrix entries in the $|T_0\rangle$ – $|T_1\rangle$ basis:

$$|T_0\rangle\langle T_0| : \frac{1 - 5p_1 - 8p_2}{6} \left[(1 - \epsilon)^5 + 5\epsilon^3(1 - \epsilon)^2 \right] + \frac{p_1}{36} \left[19 - 87\epsilon + 197\epsilon^2 - 164\epsilon^3 + 6\epsilon^4 + 32\epsilon^5 \right] + \frac{p_2}{54} \left[20 - 44\epsilon + 107\epsilon^2 - 106\epsilon^3 + 28\epsilon^4 + 8\epsilon^5 \right], \quad (10)$$

$$|T_1\rangle\langle T_1| : \frac{1 - 5p_1 - 8p_2}{6} \left[\epsilon^5 + 5\epsilon^2(1 - \epsilon)^3 \right] + \frac{p_1}{36} \left[3 + \epsilon + 61\epsilon^2 - 180\epsilon^3 + 166\epsilon^4 - 32\epsilon^5 \right] + \frac{p_2}{54} \left[13 - 4\epsilon + 37\epsilon^2 - 86\epsilon^3 + 68\epsilon^4 - 8\epsilon^5 \right], \quad (11)$$

$$|T_0\rangle\langle T_1| : \frac{p_2(1 + i)(-1 + 2\epsilon)}{432} \left[(-2 + 3i) - (1 + 5i)\sqrt{3} + ((-2 - 9i) + (3 + 10i)\sqrt{3})\epsilon + ((6 + 9i) - (3 + 6i)\sqrt{3})\epsilon^2 + ((-8 - 6i) + (2 - 8i)\sqrt{3})\epsilon^3 + (4 + 4i\sqrt{3})\epsilon^4 \right], \quad (12)$$

where $\epsilon = 1 - \langle T_0 | \rho_{\text{in}} | T_0 \rangle$ is the error of the initial state. Note that depolarizing noise in the decoding procedure will introduce off diagonal terms in the $|T_0\rangle$ – $|T_1\rangle$ basis, that is, will produce an output state that deviates away from the magic state axis. However, such matrix elements will be on the order of p_2 , which for low levels of noise will be negligible compared to the strength of the coefficients along the diagonal terms. Therefore, we may assume that upon iterating the distillation protocol, the input states will always remain along the magic state axis. This assumption is also well motivated from the results of Section II that show that slight deviations off the magic state axis will not affect the convergence of the distillation scheme. For small values of ϵ , the output error ϵ_{out} , the normalized coefficient of the $|T_1\rangle\langle T_1|$ term, can be approximated to be $5\epsilon^2 + p_1/2 + 13p_2/9$. Thus in the limit of infinite iterations of the distillation protocol the error rate will be linear in the terms p_1 and p_2 .

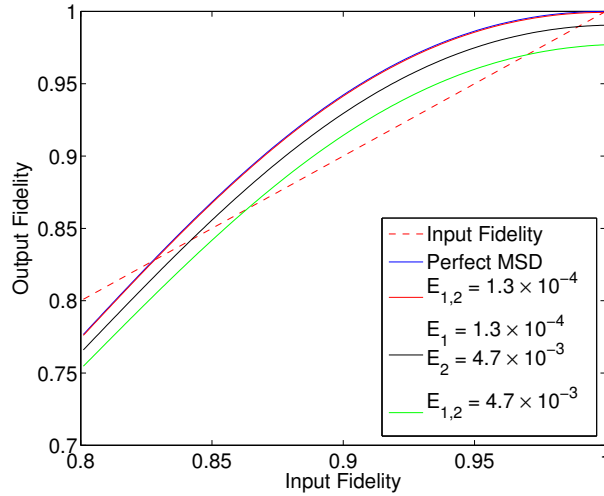


FIG. 6: Output fidelities of the magic state distillation protocol as a function of the input fidelity of the five input states. The uppermost line indicates the noiseless magic state distillation, while the three other lines show a decrease of the output fidelity caused by an increase in the values of E_1 and E_2 . As the error strength increases, the class of states for which the distillation protocol is beneficial decreases. Notice that once errors are introduced one can no longer increase the fidelity to be arbitrarily close to 1 for repeated applications of the protocol.

Motivated by the average one and two-qubit gate errors in a recent benchmarking experiment in nuclear NMR [25], we have chosen different values for the noise parameters p_1 and p_2 . The graph in Figure 6 shows the output fidelity

of the magic state distillation protocol for different error strengths of the one and two-qubit gates. The results for error strengths given by the benchmarking NMR results, $E_1 = 1.3 \times 10^{-4}$ and $E_2 = 4.7 \times 10^{-3}$, are given by the black curve. In order to obtain an appreciation of the decrease in output fidelity cause by the strength of both of these errors, we have also plotted the output fidelity corresponding to the case when the one and two-qubit errors have the same strength, namely 1.3×10^{-4} and 4.7×10^{-3} . Note that when both errors are on the order of 10^{-4} the decrease in fidelity is minimal, this leads one to conclude that the decrease in fidelity in the black curve is caused mostly by the larger error on the two-qubit gates. Finally, note that for the case of choosing an error model based on the results from Ref. [25], the new threshold for the minimal input fidelity is 0.842, which is larger than the theoretical noiseless threshold of 0.8273, and the maximum output fidelity that can be reached through repeated applications of the distillation protocol is 0.9895 with respect to the magic state.

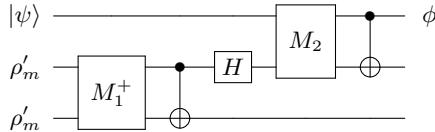


FIG. 7: Gate sequence in order to apply the $\Lambda_{\pi/12}$ gate to an arbitrary input state $|\psi\rangle$ [1]. The states ρ_m are the magic states obtained from the magic state distillation protocol. M_1 and M_2 are measurements of the observable $Z \otimes Z$. In the case of M_1^+ , one needs to post-select upon obtaining the “+1” outcome. Depending of the result of the measurement M_2 , either a $\Lambda_{\pi/12}$ or $\Lambda_{-\pi/12}$ gate is applied and the output state is ϕ . The remaining two qubits are discarded.

Due to the errors on the applied Clifford gates in the decoding procedure, distilling states arbitrarily close to the magic state is no longer possible, thus the distilled output state will have the form,

$$\rho'_m = (1 - \epsilon')|T_0\rangle\langle T_0| + \epsilon'|T_1\rangle\langle T_1|, \quad (13)$$

and the value of the error ϵ' will be fixed away from 0. However, the application of a universal non-Clifford gate is still possible with such a state, albeit with a certain level of noise. Following the procedure laid out in Ref. [1], shown in Figure 7, one can achieve the application of the gate $\Lambda_{\pi/12}$ to an arbitrary state $|\psi\rangle = a|0\rangle + b|1\rangle$ by using multiple copies of the noisy magic state ρ'_m . The fidelity of the applied gate is

$$F = 1 - \frac{12\epsilon'|a|^2(1 - |a|)^2}{3 + (1 - 2\epsilon')^2} \geq 1 - \epsilon',$$

where ϵ' is the error rate of the distilled magic state given in Equation 13. After multiple iterations of the distillation protocol, if the error of initial state was below the threshold for state distillation, the error rate ϵ' will be approximately $p_1/2 + 13p_2/9$. Thus, in order to apply the universal gate with fidelity above $1 - \epsilon'$, one would typically have to apply a final fault-tolerant iteration in order to reduce the error rates of the original Clifford gates to a smaller logical error, as will be discussed in Section III B.

B. Comparing faulty magic state distillation and fault-tolerant magic state distillation

Fault-tolerant quantum computation is a method to reduce the error rate of a given quantum operation using states and gates encoded into a higher-dimensional Hilbert space. Such encodings, combined with the ability of projective measurement and post-selection, provide a means to increase the fidelity of our quantum gates at the expense of using additional qubits and quantum gates to perform the desired encodings of states and encoded operations [5–10].

Consider a scenario where we are presented with a physical device that is not restricted in the number of qubits at its disposal but limited in the number of gates that can be applied coherently. We could then envision two different methods for applying the magic state distillation procedure, either by applying the protocol with the faulty gates at our disposal or first performing a fault-tolerant encoding of the qubits and then performing the distillation protocol with less noisy encoded logical quantum gates. The faulty distillation protocol, as described and analyzed in Section III A, suffers from a reduced convergence rate and a fundamental limit in the ability to distill a magic state with high fidelity. Thus using a fault-tolerance encoding of the Clifford gates appears to be advantageous. However the cost associated with encoding each of the Clifford gates in a fault-tolerant encoding outweighs the savings one would obtain in the number of Clifford gates one would need to apply at lower noise rates in the faulty magic state distillation scheme.

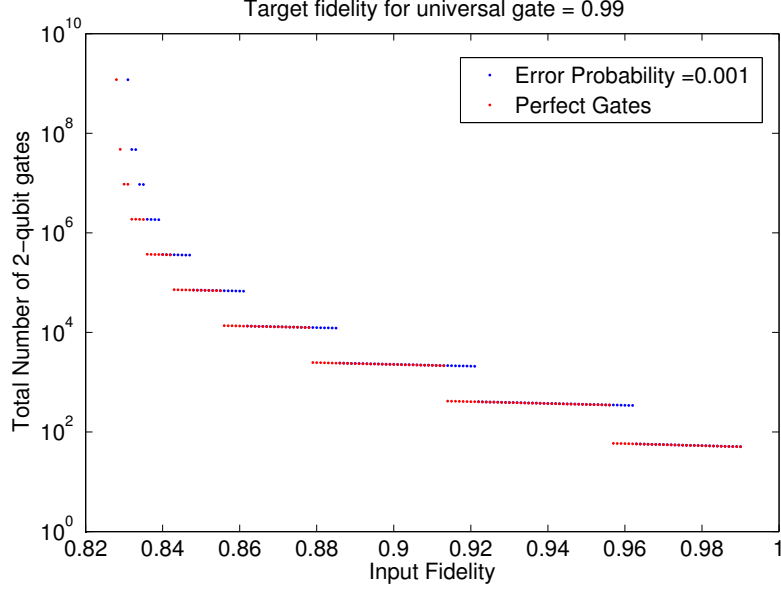


FIG. 8: Total number of two-qubit gates as a function of input fidelity for the state distillation scheme to achieve an output fidelity of 0.99 for the magic state. The red dots denote the number of gates required for perfect gates and can be thought of as a lower bound in the number of fault-tolerant encoded logical gates that would be required. The blue dots represent the total number of gates for the faulty distillation protocol, where the gate error probability E_1, E_2 is 0.001. Note that the y -axis in this figure is plotted logarithmically in order to capture that the increase in the number of gates for each iteration of the distillation scheme scales approximately by a factor of 5.

Figure 8 plots the total number of two-qubit gates that would be required for either scheme to achieve a desired final fidelity for the magic state. The red dots plotted in Figure 8 are the number of gates in the ideal setting with no noise, this can be thought of as a lower bound of the number of *logical* fault-tolerant gates that would need to be applied for any fault-tolerant encoding. As expected, the number of gates exhibit a step function behaviour, characteristic of the number of iterations of the distillation subroutine that would be required to achieve the final fidelity target. The smearing of the steps is due to the probability of measuring the trivial syndrome, which is reduced for initial states with lower fidelity. One notices that the number of iterations for the fault-tolerant magic state distillation is lower than that of the faulty procedure, this is due to the fact that noise in faulty Clifford gates can decrease the convergent rate for states along the magic state axis and will lower the probability of measuring the trivial syndrome. Thus if one were to be presented with faulty Clifford gates and were to encode the states such that the error rate of the encoded operations were negligible, then the total number of encoded *logical* gates required would be lower than the number of faulty Clifford gates to increase the fidelity of the magic state to a desired level. However, in the regime of typical one and two-qubit gate errors, 10^{-3} – 10^{-2} , the fault-tolerant encodings whose threshold rates are above these levels typically use on the order of 100–1000 two-qubit physical gates per encoded logical gate [26]. As such, in order for a fault-tolerant encoded scheme to show an improvement in the total number of gates required for the full distillation protocol, the encoded scheme would have to use on the order of 100–1000 times less logical gates than that of the faulty distillation scheme. Consider the results from Figure 8, which compares the number of faulty gates, with error probability 10^{-3} , and the ideal case of encoded gates with no error. The reduction in the number of gates with no error compared to the faulty distillation scheme is greatest at the jumps in the number of gates that occur due to an increased number of iterations required. As such, the faulty distillation scheme may require an additional iteration of the distillation scheme, which is just repeating 5 times the previous level of distillation, leading to an overall gate cost on the order of 5 times more gates. However, as mentioned above, in order to reduce the error rate of the encoded states to be below that of the original error rate, each encoded *logical* gate would require on the order of 100–1000 *physical* gates. Therefore, while the number of logical gates is lower than the number of faulty gates required, the cost of each of the logical gates would drive the overall number of gates required in a fault-tolerant scheme to be much larger. We thus conclude that using faulty Clifford gates is more efficient for magic state distillation at noise levels comparable to those in current quantum information implementations. This analysis considers the number of two-qubit gates that would be required in the distillation protocol for either scheme as typical numbers of two-qubit gates have been studied extensively in past works [26], however such an analysis could be extended to one-qubit gates and the authors believe that the behaviour will be equivalent.

Finally, one should note that fault-tolerance would be required if the errors in the one and two-qubit gates were too high, preventing the distilled state to achieve the desired target fidelity for a prepared magic state after multiple rounds of the distillation. This is due to the error of the distilled state always being approximately bounded from below by $p_1/2 + 13p_2/9$, where p_1 and p_2 are the error probabilities of the one and two-qubit gates, respectively. Therefore, if one requires to reduce the error of the distilled magic state, and subsequently the applied universal gate using the magic states, to be below some small target threshold, one would need to reduce the error rates of the one and two-qubit gates through a method such as fault-tolerance. However, this step would only have to be applied in the final iterations of the protocol in order to get the distilled state over the fidelity hump posed by the errors in the Clifford gates at one's disposal.

IV. CONCLUSION

We have both analytically and numerically analyzed the evolution of quantum states under the five-qubit magic state distillation protocol without the presence of the dephasing transformation for perfect Clifford gates. We have characterized which input states converge to the magic state under repeated application of the distillation subroutine. Section II provides results showing that not all states on a plane with fidelity over the threshold converge to magic state after multiple iterations of subroutine without dephasing. However, some quantum states which are undistillable become distillable with the addition of the dephasing transformation for high enough fidelity, as was shown by Bravyi and Kitaev [1]. Therefore, the fact that some quantum states contribute to universal quantum computation is dependent on the ability to access the dephasing transformation. Thus, on a more fundamental level, one may ask what is the role of the dephasing transformation in universal quantum computation? Also noted in Section II, there are states below the theoretical threshold derived for states along the magic state axis, which converge to the magic state after many repetitions of the distillation routine, and as such if one had access to such states off the magic state axis, one could lower the threshold for distillation to 0.825. An interesting direction for future research would be to investigate whether other distillation schemes [14–18] are also robust to such noise perturbations.

Additionally, we studied the effect of noisy Clifford gates on the output fidelity, and the error rate in the universal gate induced by noisy distillation of magic states. Fortunately, based on the average error strength given in a recent benchmarking experiment [25], the state after absorbing all the noise in the Clifford gates is not far away from the pure magic state. This means that magic state distillation is robust against the typical noise levels in current experimental implementations. Nevertheless, due to the errors on the applied Clifford gates in the distillation procedure, one cannot produce an output state arbitrarily close to the magic state. In order to further reduce the error on the magic state, and the subsequent universal gate application, one would need to introduce a fault-tolerant encoding of the Clifford gates to reduce their noise. However, it is only at this point that a fault-tolerant encoding would be beneficial. We showed in Section III B that although introducing a fault-tolerant encoding at the beginning of the distillation protocol may seem appealing due to the increased convergence rate of the protocol, it is highly inefficient due to the number of gates that are typically required in any fault-tolerant scheme to reduce the error rate of current implementations of quantum information processors [26]. Since fault-tolerance would not be required to perform magic state distillation, other than in the final iterations of the protocol in order to apply a universal gate with very high fidelity, this work confirms that multiple rounds of faulty magic state distillation is a realistic scheme for future small-scale quantum information processors.

V. ACKNOWLEDGEMENTS

Author Contributions – All authors contributed to the development and discussion of the results in this work. T.J. and Y.Y. co-wrote the manuscript.

T.J. acknowledges the support of NSERC through the Alexander Graham Bell CGS–M scholarship. Y.Y. acknowledges the support of the China Scholarship Council. B.H. acknowledges the support of NSERC through the NSERC Undergraduate Student Research Award (USRA). The authors would like to thank Joseph Emerson and Ben Reichardt for stimulating insightful discussion throughout this project. This work was additionally funded through QuantumWorks, CIFAR, and Industry Canada.

[1] S. Bravyi and A. Kitaev, “Universal quantum computation with ideal Clifford gates and noisy ancillas,” *Phys. Rev. A*, vol. 71, p. 022316, 2005.

- [2] R. Landauer, “Irreversibility and heat generation in the computing process,” *IBM Journal of Research and Development*, vol. 5, pp. 183–191, 1961.
- [3] P. W. Shor, “Scheme for reducing decoherence in quantum computer memory,” *Phys. Rev. A*, vol. 52, p. 2493, 1995.
- [4] A. W. Steane, “Error correcting codes in quantum theory,” *Phys. Rev. Lett.*, vol. 77, p. 793, 1996.
- [5] P. Shor, “Fault-tolerant quantum computation,” *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on*, pp. 56–65, oct 1996.
- [6] D. Gottesman, “Stabilizer Codes and Quantum Error Correction,” *Ph.D. thesis, California Institute of Technology, Pasadena, CA*, 1997.
- [7] E. Knill, R. Laflamme, and W. H. Zurek, “Resilient quantum computation,” *Science*, vol. 279, pp. 342–345, 1998.
- [8] J. Preskill, “Reliable quantum computers,” *Proc. R. Soc. London A*, vol. 454, pp. 385–410, 1998.
- [9] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, “Limitations of noisy reversible computation,” *arXiv:quant-ph/9611028*, 1996.
- [10] P. Aliferis, D. Gottesman, and J. Preskill, “Quantum accuracy threshold for concatenated distance-3 codes,” *Quant. Inf. Comput.*, vol. 6, pp. 97–165, 2006.
- [11] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, England, 2000.
- [12] S. Aaronson and D. Gottesman, “Improved simulation of stabilizer circuits,” *Phys. Rev. A*, vol. 70, p. 052328, 2004.
- [13] E. Knill, “Fault-tolerant postselected quantum computation: schemes,” *arXiv:quant-ph/0402171*, 2004.
- [14] B. W. Reichardt, “Quantum universality from magic states distillation applied to CSS codes,” *Quant. Info. Proc.*, vol. 4, p. 251, 2005.
- [15] B. W. Reichardt, “Quantum universality by state distillation,” *Quant. Inf. Comput.*, vol. 9, pp. 1030–1052, 2005.
- [16] E. Campbell and D. Browne, *On the Structure of Protocols for Magic State Distillation*, vol. 5906 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2009.
- [17] E. T. Campbell and D. E. Browne, “Bound states for magic state distillation in fault-tolerant quantum computation,” *Phys. Rev. Lett.*, vol. 104, p. 030503, 2010.
- [18] H. Anwar, E. T. Campbell, and D. E. Browne, “Qutrit magic state distillation,” *arXiv:1202.2326*, 2012.
- [19] V. Veitch, C. Ferrie, and J. Emerson, “Negative Quasi-Probability Representation is a Necessary Resource for Magic State Distillation,” *arXiv:1201.1256*, 2012.
- [20] J. T. Anderson, “On the Power of Reusable Magic States,” *arXiv:1205.0289*, 2012.
- [21] A. M. Souza, J. Zhang, C. A. Ryan, and R. Laflamme, “Experimental magic state distillation for fault-tolerant quantum computing,” *Nature Comm.*, vol. 2, no. 169, 2011.
- [22] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, “Mixed state entanglement and quantum error correction,” *Phys. Rev. A*, no. 54, p. 3824, 1996.
- [23] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, “Bound states for magic state distillation in fault-tolerant quantum computation,” *Phys. Rev. Lett.*, vol. 77, no. 198, 1996.
- [24] M. Grassl, “Bounds on the minimum distance of linear codes and quantum codes.” Online available at <http://www.codetables.de>, 2007.
- [25] C. A. Ryan, M. Laforest, and R. Laflamme, “Randomized benchmarking of single and multi-qubit control in liquid-state NMR quantum information processing,” *New J. Phys.*, vol. 11, p. 013034, 2009.
- [26] A. Cross, D. P. DiVincenzo, and B. M. Terhal, “A comparative code study for quantum fault tolerance,” *Quant. Inf. Comput.*, vol. 9, pp. 541–572, 2009.